Unifying Decoherence and the Heisenberg Principle

Bas Janssens

February 1, 2008

Abstract

We exhibit three inequalities involving quantum measurement, all of which are sharp and state independent. The first inequality bounds the performance of joint measurement. The second quantifies the trade-off between the measurement quality and the disturbance caused on the measured system. Finally, the third inequality provides a sharp lower bound on the amount of decoherence in terms of the measurement quality. This gives a unified description of both the Heisenberg uncertainty principle and the collapse of the wave function.

1 Introduction

In quantum mechanics, observables are modelled by self-adjoint operators A, and states by normalized trace-class operators ρ. A state ρ induces a probability measure on an observable A. It is the objective of a *quantum measurement* to portray this probability measure as faithfully as possible.

According to the uncertainty relation $\sigma_X \sigma_Y \geq \frac{1}{2}|\mathbf{tr}(\rho[X,Y])|$, (see [He, Ke, Ro]), there is an inherent variance in the quantum state. Furthermore, quantum theory puts severe restrictions on the performance of measurement. These restrictions, which come on top of the measurement restrictions implied by the above uncertainty relation, fall into three distinct classes.

   I The impossibility of perfect *joint measurement*. It is not possible to perform a simultaneous measurement of two noncommuting observables in such a way that both measurements have perfect quality.

  II The *Heisenberg principle*, (see [He]). This states that quantum information cannot be extracted from a system without disturbing that system.

 III The *collapse of the wave function*. When information is extracted from a quantum system, a so-called decoherence is experimentally known to occur on this system.

We will see that this collapse of the wave function is a mathematical consequence of information extraction. In the process, II and III will be clearly exhibited as two sides of the same coin.

The subject of uncertainty relations in quantum measurement is already endowed with an extensive literature. For example, the Heisenberg principle and the impossibility of joint measurement are quantitatively illustrated in [AK, Oz, Is, Ha].

However, the inequalities in these papers depend on the state ρ, which somewhat limits their practical use. Indeed, the bound on the measurement quality can only be calculated if the state ρ is known, in which case there is no need for a measurement in the first place.

Our state-independent figures of merit (sections 2 and 3) will lead us quite naturally to state-independent bounds on the performance of measurement. In order to illustrate their practical

use, we will give some applications. We investigate the beamsplitter, resonance fluorescence and nondestructive qubit measurement.

In section 4, we will prove a sharp, state independent bound on the performance of jointly unbiased measurement. This generalizes the impossibility of perfect joint measurement.

In section 5, we will prove a sharp, state independent bound on the performance of a measurement in terms of the maximal disturbance that it causes. This generalizes the Heisenberg principle.

In contrast with the Heisenberg principle and its abundance of inequalities, the phenomenon of decoherence has mainly been investigated in specific examples (see e.g. [Hp, Zu, JZ]). Although there are some bounds on the remaining coherence in terms of the measurement quality (see [JM, Se]), a sharp, information-theoretic inequality does not yet appear to exist.

We will provide such an inequality in section 6, where we will prove a sharp upper bound on the amount of coherence which can survive information transfer. Not only does this generalize the collapse of the wave function, it also shows that no information can be extracted if all coherence is left perfectly intact. It is therefore a *unified* description of both the Heisenberg principle and the collapse of the wave function.

2   Information Transfer

In quantum mechanics, a system is described by a von Neumann algebra $\mathcal{A}$ of bounded operators on a Hilbert space $\mathcal{H}$. (Usually the algebra $B(\mathcal{H})$ of all bounded operators.) Its state space is formed by the normalized density matrices $\mathcal{S}(\mathcal{A}) = \{\rho \in \mathcal{A} \,;\, \rho \geq 0, \, \mathbf{tr}(\rho) = 1\}$. With the system in state $\rho \in \mathcal{S}(\mathcal{A})$, observation of a (Hermitean) observable $A \in \mathcal{A}$ is postulated to yield the average value $\mathbf{tr}(\rho A)$.

Definition. Let $\mathcal{A}$ and $\mathcal{B}$ be von Neumann algebras. A map $T : \mathcal{B} \to \mathcal{A}$ is called *Completely Positive* (or CP for short) if it is linear, normalized (i.e. $T(\mathbf{1}) = \mathbf{1}$), positive (i.e. $T(X^{\dagger}X) \geq 0$ for all $X \in \mathcal{B}$) and if moreover the extension $Id_n \otimes T : M_n \otimes \mathcal{B} \to M_n \otimes \mathcal{A}$ is positive for all $n \in \mathbb{N}$, where M_n is the algebra of complex $n \times n$-matrices. In this article, we will require CP-maps to be weakly continuous unless specified otherwise.

Its dual $T^* : \mathcal{S}(\mathcal{A}) \to \mathcal{S}(\mathcal{B})$, defined by the requirement $\mathbf{tr}(T^*(\rho)X) = \mathbf{tr}(\rho T(X)) \, \forall \, X \in \mathcal{B}$, has a direct physical interpretation as an operation between quantum systems. First of all, due to positivity and normalization of T, each state $\rho \in \mathcal{S}(\mathcal{A})$ is again mapped to a state $T^*(\rho) \in \mathcal{S}(\mathcal{B})$. Secondly, linearity implies that T^* satisfies $pT^*(\rho_1) + (1-p)T^*(\rho_2) = T^*(p\rho_1 + (1-p)\rho_2)$ for all $p \in [0, 1]$, $\rho_1, \rho_2 \in \mathcal{S}(\mathcal{A})$. This expresses the *stochastic equivalence principle*: a system which is in state ρ_1 with probability p and in state ρ_2 with probability $(1 - p)$ cannot be distinguished from a system in state $p\rho_1 + (1 - p)\rho_2$. Finally, it is possible to extend the systems $\mathcal{A}$ and $\mathcal{B}$ under consideration with another system M_n, on which the operation acts trivially. Due to complete positivity, states in $\mathcal{S}(M_n \otimes \mathcal{A})$ are once again mapped to states in $\mathcal{S}(M_n \otimes \mathcal{B})$. Incidentally, any CP-map T automatically satisfies $T(X^{\dagger}) = T(X)^{\dagger}$ and $\|T(X)\| \leq \|X\|$ for all $X \in \mathcal{B}$.

2.1   General, Unbiased and Perfect Information Transfer

Suppose that we are interested in the distribution of the observable $A \in \mathcal{A}$, with the system $\mathcal{A}$ in some unknown state ρ. We perform the operation $T^* : \mathcal{S}(\mathcal{A}) \to \mathcal{S}(\mathcal{B})$, and then observe the 'pointer' B in $\mathcal{B}$ in order to obtain information on A. One may (see [Ha]) take the position that *any* CP-map $T : \mathcal{B} \to \mathcal{A}$ is an information transfer from *any* observable $A \in \mathcal{A}$ to *any* pointer $B \in \mathcal{B}$. The following is a figure of demerit for the quality of such an information transfer.

Definition. Let $T : \mathcal{B} \to \mathcal{A}$ be a CP-map. Its *measurement infidelity* δ in transferring information from A to the pointer B is defined as $\delta := \sup_S \|\mathbf{1}_S(A) - T(\mathbf{1}_S(B))\|$, where S runs over the Borel subsets of $\mathbb{R}$.

It measures how accurately probability distributions on the measured observable A are copied to the pointer B.

The initial state ρ defines a probability distribution $\mathbb{P}_i$ on the spectrum of A by $\mathbb{P}_i(S) := \mathbf{tr}(\rho\mathbf{1}_S(A))$, where $\mathbf{1}_S(A)$ denotes the spectral projection of A associated to the set S. Similarly, the final state $T^*(\rho)$ defines a probability distribution $\mathbb{P}_f$ on the spectrum of B. δ is now the maximum distance between $\mathbb{P}_i$ and $\mathbb{P}_f$, where the maximum is taken over all initial states ρ. That is, $\delta = \sup_\rho D(\mathbb{P}_i, \mathbb{P}_f)$.

The *trace distance* (a.k.a. variational distance or Kolmogorov distance) is defined as $D(\mathbb{P}_f, \mathbb{P}_i) := \sup_S\{|\mathbb{P}_i(S) - \mathbb{P}_f(S)|\}$, the difference between the probability that the event S occurs in the distribution $\mathbb{P}_i$ and the probability that it occurs in the distribution $\mathbb{P}_f$, for the worst case Borel set S. Writing out this definition, we see that indeed $\sup_\rho D(\mathbb{P}_i, \mathbb{P}_f) = \sup_{\rho,S} |\mathbf{tr}(\rho\mathbf{1}_S(A)) - \mathbf{tr}(\rho T(\mathbf{1}_S(B)))| = \sup_S \|\mathbf{1}_S(A) - T(\mathbf{1}_S(B))\| = \delta$. The measurement infidelity δ is precisely the worst case difference between input and output probabilities.

In this article, we will devote considerable attention to the class of unbiased information transfers.

Definition. A CP-map $T : \mathcal{B} \to \mathcal{A}$ is called an *unbiased information transfer* from the Hermitean observable $A \in \mathcal{A}$ to a Hermitean $B \in \mathcal{B}$ if $T(B) = A$.

Recall that we are interested in the distribution of A, with the system $\mathcal{A}$ in some unknown state ρ. We perform the operation $T^* : \mathcal{S}(\mathcal{A}) \to \mathcal{S}(\mathcal{B})$, and then observe the 'pointer' B in $\mathcal{B}$. Since $\mathbf{tr}(T^*(\rho)B) = \mathbf{tr}(\rho T(B))$ by definition of the dual, and $\mathbf{tr}(\rho T(B)) = \mathbf{tr}(\rho A)$ by definition of unbiased information transfer, the expectation value of B in the final state $T^*(\rho)$ is the same as that of A in the initial state ρ. We conclude that the expectation of A was transferred to B.

Definition. An information transfer $T : \mathcal{B} \to \mathcal{A}$ from $A \in \mathcal{A}$ to $B \in \mathcal{B}$ is called *perfect* if $T(B) = A$ and if the restriction of T to B'', the von Neumann algebra generated by B, is a *-homomorphism $B'' \to A''$.

The entire probability distribution of A is then transferred to B, rather than merely its average value. Indeed, for all moments $\mathbf{tr}(\rho A^n)$, we have $\mathbf{tr}(T^*(\rho)B^n) = \mathbf{tr}(\rho T(B^n)) = \mathbf{tr}(\rho T(B)^n) = \mathbf{tr}(\rho A^n)$. Everything there is to know about A in the initial state ρ can be obtained by observing the 'pointer' B in the final state $T^*(\rho)$. Note that the transfer is perfect if and only if $\delta = 0$.

Schematically, we have {General information transfers} $\supset$ {Unbiased information transfers} $\supset$ {Perfect information transfers}.

2.2   Example: von Neumann Qubit Measurement

Let $\Omega := \{+1, -1\}$. Denote by $\mathcal{C}(\Omega)$ the (commutative) algebra of $\mathbb{C}$-valued random variables on Ω. A state on $\mathcal{C}(\Omega)$ is precisely a probability distribution $\mathbb{P}$ on Ω, and $\mathbf{tr}(\mathbb{P}f)$ should be read as $\mathbb{E}(f)$. Define the probability distributions $\mathbb{P}_\pm$ to assign probability 1 to ± 1.

The Von Neumann-measurement $T : M_2 \otimes \mathcal{C}(\Omega) \to M_2$ is defined as $T(X \otimes f) := f(+1)P_+XP_+ + f(-1)P_-XP_-$, with $P_+ = |\uparrow\rangle\langle\uparrow|$ and $P_- = |\downarrow\rangle\langle\downarrow|$. Then $T^* : \mathcal{S}(M_2) \to \mathcal{S}(M_2) \otimes \mathcal{S}(\mathcal{C}(\Omega))$ is given by $T^*(\rho) = \mathbf{tr}(\rho P_+)|\uparrow\rangle\langle\uparrow| \otimes \mathbb{P}_+ + \mathbf{tr}(\rho P_-)|\downarrow\rangle\langle\downarrow| \otimes \mathbb{P}_-$.

In words: with probability $\mathbf{tr}(\rho P_+)$, the output $+1$ occurs and the qubit is left in state $|\uparrow\rangle$. With probability $\mathbf{tr}(\rho P_-)$, the output -1 occurs, leaving the qubit in state $|\downarrow\rangle$. The Von

Neumann-measurement T is a perfect (and thus unbiased) information transfer from $\sigma_z \in M_2$ to $\mathbf{1} \otimes (\delta_{+1} - \delta_{-1}) \in M_2 \otimes \mathcal{C}(\Omega)$.

Quantum measurements are often (e.g. [Ho, Ha]) modelled by Positive Operator Valued Measures or POVM's. From our CP-map T, we may distill the POVM $\mu : \Omega \to M_2$ by $\mu(\omega) := T(\mathbf{1} \otimes \delta_\omega)$, i.e. $\mu(+1) = P_+$ and $\mu(-1) = P_-$. This procedure is fully general: *any* CP-map gives rise to a POVM on a suitable Σ-algebra.

A CP-map can thus be seen as an extension of a POVM that keeps track of the system output as well as the measurement output. Since we will be interested in disturbance of the system, it is imperative that we consider the full CP-map rather than merely its POVM.

3 Maximal Added Variance

For unbiased information transfer, there exists a figure of demerit more attractive than δ. Consider the variance $\mathbf{Var}(B, T^*(\rho))$ of the output. (The variance is defined as $\mathbf{Var}(X, \rho) := \mathbf{tr}(\rho X^\dagger X) - \mathbf{tr}(\rho X)^* \mathbf{tr}(\rho X)$.) The output variance can be split in two parts. One part $\mathbf{Var}(A, \rho)$ is the variance of the input, which is intrinsic to the quantum state ρ. The other part $\mathbf{Var}(B, T^*(\rho))$ $- \mathbf{Var}(A, \rho) \geq 0$ is *added* by the measurement procedure. This second part determines how well the measurement performs.

The maximal added variance (where the maximum is taken over the input states ρ) will be our figure of demerit. For example, perfect information transfer from A to B satisfies $\mathbf{Var}(B, T^*(\rho)) = \mathbf{Var}(A, \rho)$, so that the maximal added variance is 0. There is uncertainty in the measurement outcome, but all uncertainty 'comes from' the quantum state, and none is added by the measurement procedure.

Definition. The *maximal added variance* of an unbiased information transfer T is defined as

$$\Sigma^2 := \sup_{\rho \in \mathcal{S}(\mathcal{A})} \mathbf{Var}(B, T^*(\rho)) - \mathbf{Var}(A, \rho) \,.$$

It is straightforward to verify $\Sigma^2 = \|T(B^\dagger B) - T(B)^\dagger T(B)\|$. This inspires the following definition.

Definition. Let $T : \mathcal{B} \to \mathcal{A}$ be a CP-map. We define the operator-valued sesquilinear form $(\cdot, \cdot) : \mathcal{B} \times \mathcal{B} \to \mathcal{A}$ by

$$(X, Y) := T(X^\dagger Y) - T(X)^\dagger T(Y) \,.$$

It satisfies $(X, Y)^\dagger = (Y, X)$ and is positive semi-definite: $(B, B) \geq 0$ for all $B \in \mathcal{B}$. This 'length' has the physical interpretation $\|(B, B)\| = \Sigma^2$, and there is even a Cauchy-Schwarz inequality:

Lemma 1 (Cauchy-Schwarz) *Let* $T : \mathcal{B} \to \mathcal{A}$ *be a CP-map, and* $(X, Y) := T(X^\dagger Y) - T(X)^\dagger T(Y)$. *Then for all* $X, Y \in \mathcal{B}$:

$$(X, Y)(Y, X) \leq \|(Y, Y)\|(X, X).$$

Proof: By the Stinespring-theorem (see [Ta]), we may assume without loss of generality that T is of the form $T(X) = V^\dagger X V$ for some contraction V. Writing this out, we obtain $(X, Y) = V^\dagger X^\dagger (\mathbf{1} - VV^\dagger) Y V$. Defining $g(X) := \sqrt{\mathbf{1} - VV^\dagger} X V$, we write $(X, Y) = g(X)^\dagger g(Y)$. Thus $(X, Y)(Y, X) = g(X)^\dagger g(Y) g(Y)^\dagger g(X) \leq \|g(Y)\|^2 g(X)^\dagger g(X) = \|(Y, Y)\|(X, X)$. $\qquad\square$

If an information transfer is perfect, then of course $\Sigma^2 = \|(B,B)\| = 0$. (No variance is added.) We will now show that the converse also holds: if $\Sigma^2 = 0$, then T is a *-homomorphism on B''. (Compare this with the fact that probability distributions of 0 variance are concentrated in a single point.)

Theorem 2 *Let $T : \mathcal{B} \to \mathcal{A}$ be a CP-map, let $B \in \mathcal{B}$ be Hermitean. Then among*

1. *$(B,B) = 0$.*
2. *The restriction of T to B'', the von Neumann algebra generated by B, is a *-homomorphism $B'' \to T(B)''$.*
3. *$(f(B), f(B)) = 0$ for all measurable functions f on the spectrum of B.*
4. *T maps the relative commutant $B' = \{X \in \mathcal{A}; [X,B] = 0\}$ into $T(B)'$.*

the following relations hold: $(1) \Leftrightarrow (2) \Leftrightarrow (3) \Rightarrow (4)$.

Proof: First $(1) \Rightarrow (2)$. By Cauchy-Schwarz (lemma 1), we have $T(B^n) - T(B)T(B^{n-1}) \leq \|(B,B)\|(B^{n-1}, B^{n-1}) = 0$. By induction, $T(B^n) = T(B)^n$, and by linearity $T(f(B)) = f(T(B))$ for all polynomials f. Thus T is a *-homomorphism from the algebra of polynomials on the spectrum of B to that on $T(B)$. Since T is weakly continuous, this statement extends to the algebras of measurable functions on the spectra of B and $T(B)$, isomorphic to B'' and $T(B)''$ respectively. For $(2) \Rightarrow (3)$, note that $T(f(B)^2) = T(f(B))^2$. For $(3) \Rightarrow (1)$, take $f(x) = x$. Finally we prove $(1) \Rightarrow (4)$: suppose that $[A,B] = 0$. Then $[T(B), T(A)] = T([A,B]) - [T(A), T(B)] = (A^\dagger, B) - (B^\dagger, A)$. ($B$ is Hermitean.) By Cauchy-Schwarz, the last term equals zero if (B,B) does. $\qquad\square$

We see that the maximal added variance Σ^2 equals 0 if and only if T is a perfect information transfer. We shall take Σ to parametrize the imperfection of unbiased information transfer.

4   Joint Measurement

In a jointly unbiased measurement, information on two observables A and $\tilde{A}$ is transferred to two *commuting* pointers B and $\tilde{B}$. If A and $\tilde{A}$ do not commute, then it is not possible for both information transfers to be perfect. (See [Ne], [We2].) Indeed, the degree of imperfection is determined by the amount of noncommutativity:

Theorem 3 *Let $T : \mathcal{B} \to \mathcal{A}$ be a CP-map, let $B, \tilde{B}$ be commuting Hermitean observables in $\mathcal{B}$, and define $A := T(B)$, $\tilde{A} := T(\tilde{B})$, $\Sigma_B^2 := \|(B,B)\|$ and $\Sigma_{\tilde{B}}^2 := \|(\tilde{B}, \tilde{B})\|$. Then*

$$\Sigma_B \Sigma_{\tilde{B}} \geq \tfrac{1}{2} \|[A, \tilde{A}]\|. \tag{1}$$

Proof: Since $[B, \tilde{B}] = 0$, we have $[\tilde{A}, A] = T([B, \tilde{B}]) - [T(B), T(\tilde{B})] = (B, \tilde{B}) - (\tilde{B}, B)$. By Cauchy-Schwarz, the latter is at most $2\Sigma_B \Sigma_{\tilde{B}}$ in norm. $\qquad\square$

We now show that this bound is sharp in the sense that for all $S, \tilde{S} > 0$, there exist $T, B, \tilde{B}$ such that (1) attains equality with $\Sigma_B = S$, $\Sigma_{\tilde{B}} = \tilde{S}$.

4.1   Application: the Beamsplitter as a Joint Measurement

A *beamsplitter* is a device which takes two beams of light as input. A certain fraction of each incident beam is refracted and the rest is reflected, in such a way that the refracted part of the first beam coincides with the reflected part of the second and vice versa. We will show that the beamsplitter serves as an optimal joint unbiased measurement.
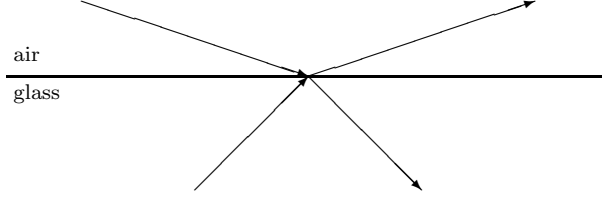
Fig. 1: Beamsplitter.

In cavity QED, a single mode in the field is described by a Hilbert space $\mathcal{H}$ of a harmonic oscillator, with creation and annihilation operators $a^\dagger$ and a satisfying $[a, a^\dagger] = 1$, as well as $x = \frac{a+a^\dagger}{\sqrt{2}}$ and $p = \frac{a-a^\dagger}{\sqrt{2}i}$. The coherent states $|\alpha\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^\infty \frac{\alpha^n}{\sqrt{n!}} |n\rangle$ are dense in $\mathcal{H}$, and satisfy $a|\alpha\rangle = \alpha|\alpha\rangle$.

Quantummechanically, a beamsplitter is described by the unitary operator U on $\mathcal{H} \otimes \mathcal{H}$, given by $U = \exp(\theta(a^\dagger \otimes a - a \otimes a^\dagger))$. In terms of the coherent vectors, we have $U|\alpha\rangle \otimes |\beta\rangle = |\alpha\cos(\theta) + \beta\sin(\theta)\rangle \otimes |-\alpha\sin(\theta) + \beta\cos(\theta)\rangle$. Note that $U^\dagger a \otimes \mathbf{1} U = \cos(\theta)a \otimes \mathbf{1} + \sin(\theta)\mathbf{1} \otimes a$ and that $U^\dagger \mathbf{1} \otimes aU = -\sin(\theta)a \otimes \mathbf{1} + \cos(\theta)\mathbf{1} \otimes a$. (This can be seen by sandwiching both sides between coherent vectors.) Since the map $Y \mapsto U^\dagger Y U$ respects $+$, $\cdot$ and †, we readily calculate

$$
\begin{aligned}
U^\dagger x \otimes \mathbf{1} U &= \cos(\theta)x \otimes \mathbf{1} + \sin(\theta)\mathbf{1} \otimes x, \\
U^\dagger x^2 \otimes \mathbf{1} U &= \cos^2(\theta)x^2 \otimes \mathbf{1} + 2\sin(\theta)\cos(\theta)x \otimes x + \sin^2(\theta)\mathbf{1} \otimes x^2, \\
U^\dagger \mathbf{1} \otimes pU &= -\sin(\theta)p \otimes \mathbf{1} + \cos(\theta)\mathbf{1} \otimes p, \\
U^\dagger \mathbf{1} \otimes p^2U &= \sin^2(\theta)p^2 \otimes \mathbf{1} - 2\cos(\theta)\sin(\theta)p \otimes p + \cos^2(\theta)\mathbf{1} \otimes p^2.
\end{aligned}
$$

We are now interested in the map $\rho \mapsto U\rho \otimes |0\rangle\langle 0|U^\dagger$, from $\mathcal{S}(\mathcal{H})$ to $\mathcal{S}(\mathcal{H}) \otimes \mathcal{S}(\mathcal{H})$. In other words, we feed the beamsplitter only one beam of light in a state ρ, the other input being the vacuum. The dual of this is the CP-map $T : \mathcal{B}(\mathcal{H}) \otimes \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ defined by $T(Y) := id \otimes \phi_0(U^\dagger Y U)$, with ϕ_0 the vacuum state $\phi_0(X) = \langle 0|X|0\rangle$.

Take $B = \cos^{-1}(\theta)x \otimes \mathbf{1}$ for instance. Then $T(B) = x\langle 0|\mathbf{1}|0\rangle + \tan(\theta)\mathbf{1}\langle 0|x|0\rangle = x$. Similarly, with $\tilde{B} = -\sin^{-1}(\theta)\mathbf{1} \otimes p$, we have $T(\tilde{B}) = p$. Apparently, splitting a beam of light in two parts, measuring $x \otimes \mathbf{1}$ in the first beam and $\mathbf{1} \otimes p$ in the second, and then compensating for the loss of intensity provides a *simultaneous unbiased measurement* of x and p in the original beam. Since $[x, p] = i$, we must[1] have $\Sigma_B \Sigma_{\tilde{B}} \geq \frac{1}{2}$.

We now calculate Σ_B and $\Sigma_{\tilde{B}}$ explicitly. From $\langle 0|x^2|0\rangle = \frac{1}{2}$, we see that $T(B^2) = x^2 + \frac{1}{2}\tan^2(\theta)\mathbf{1}$. Thus $\Sigma_B^2 = \|(B, B)\| = \frac{1}{2}\tan^2(\theta)$. Similarly $\Sigma_{\tilde{B}}^2 = \frac{1}{2}\tan^{-2}(\theta)$. We see that $\Sigma_B \Sigma_{\tilde{B}} = \frac{1}{2}$, so that the beamsplitter is indeed an optimal jointly unbiased measurement.

By scaling B, optimal joint measurements can be found for arbitrary values of Σ_B and $\Sigma_{\tilde{B}}$, which shows the bound in theorem 3 to be sharp. It may therefore be used to evaluate joint measurement procedures. For example, it was shown in [JB] that homodyne detection of the spontaneous decay of a two-level atom constitutes a joint measurement with $\Sigma\Sigma' = 1.056$, slightly above the bound $\Sigma\Sigma' \geq 1$ provided by theorem 3.

The beamsplitter is an optimal joint measurement in the sense that it minimizes $\Sigma\Sigma'$. It also performs well with other figures of merit. For example, if the quality of joint measurement is judged by the state-dependent cost $R(T) := \mathbf{Var}(B, T^*(\rho)) + \mathbf{Var}(\tilde{B}, T^*(\rho))$, then at least for Gaussian ρ, the optimal measurement is again the above beamsplitter with $\theta = \pi/4$. (See [Ho].)

---

[1] We neglect the technical complication of x and p being unbounded operators.

5   The Heisenberg Principle

The Heisenberg Principle may be stated as follows:

If all states are left intact, no quantum-information can be extracted from a system.

This alludes to an information transfer from an initial system $\mathcal{A}$ to a final system consisting of two parts: the system $\mathcal{A}$ and an ancilla $\mathcal{B}$, containing the pointer B. We thus have an information transfer $T : \mathcal{A} \otimes \mathcal{B} \to \mathcal{A}$ from A to $\mathbf{1} \otimes B$.

An initial state $\rho \in \mathcal{S}(\mathcal{A})$ gives rise to a final state $T^*(\rho) \in \mathcal{S}(\mathcal{A} \otimes \mathcal{B})$. Restricting this final state to the system $\mathcal{A}$ (i.e. taking the partial trace over $\mathcal{B}$) yields a 'residual' state $R^*(\rho) \in \mathcal{S}(\mathcal{A})$, whereas taking the partial trace over $\mathcal{A}$ yields the final state $Q^*(\rho) \in \mathcal{S}(\mathcal{B})$ of the ancilla. We define the CP-maps $R : \mathcal{A} \to \mathcal{A}$ by $R(A) := T(A \otimes \mathbf{1})$ and $Q : \mathcal{B} \to \mathcal{A}$ by $Q(B) := T(\mathbf{1} \otimes B)$. The map R describes what happens to $\mathcal{A}$ if we forget about the ancilla $\mathcal{B}$, and Q describes the ancilla, neglecting the original system $\mathcal{A}$.

We wish to find a quantitative version of the Heisenberg principle, i.e. we want to relate the imperfection of the extracted quantum-information to the amount of state disturbance.

Definition. The *maximal disturbance* Δ of a map $R : \mathcal{A} \to \mathcal{A}$ is given by $\Delta := \sup\{\|R(P) - P\| \, ; \, P \in \mathcal{A}, P^2 = P^\dagger = P\}$.

The trace distance (or Kolmogorov distance) $D(\tau, \rho)$ is the maximal difference between the probability $\mathbf{tr}(\tau P)$ that an event P occurs in the state τ, and the probability $\mathbf{tr}(\rho P)$ that it occurs in the state ρ, for the worst case event (projection operator) P. For short, $D(\tau, \rho) := \sup_P\{|\mathbf{tr}(\tau P) - \mathbf{tr}(\rho P)|\}$. One may show that $D(\tau, \rho) = \frac{1}{2}\mathbf{tr}(|\tau - \rho|)$ (see e.g. [NC]).

Δ is now the worst case distance between the input ρ and the output $R^*(\rho)$, i.e. $\Delta = \sup\{D(\rho, R^*(\rho)); \rho \in \mathcal{S}(\mathcal{A})\}$. Indeed, $\sup_\rho\{D(\rho, R^*(\rho))\} = \sup_{\rho, P}\{\mathbf{tr}(\rho P) - \mathbf{tr}(\rho R(P))\}$, which equals $\sup_P\{\|R(P) - P\|\} = \Delta$.

5.1   Heisenberg Principle for Unbiased Information Transfer

We first turn our attention to unbiased information transfer. The imperfection of the information is then captured in the maximal added variance Σ^2.

The Heisenberg principle only holds for quantum-information. Classical observables are contained in the centre $\mathcal{Z} = \{A \in \mathcal{A} \, ; \, [A, X] = 0 \; \forall X \in \mathcal{A}\}$, whereas quantum observables are not. The degree in which an observable A is 'quantum' is given by its distance to the centre $d(A, \mathcal{Z}) = \inf_{Z \in \mathcal{Z}} \|A - Z\|$. In the following, we will take the algebra of observables to be $B(\mathcal{H})$ for some Hilbert space $\mathcal{H}$. The centre is then simply $\mathbb{C}\mathbf{1}$.

Theorem 4 *Let $T : B(\mathcal{H}) \otimes \mathcal{B} \to B(\mathcal{H})$ be a CP-map, let $B \in \mathcal{B}$ be Hermitean. Define $A := T(\mathbf{1} \otimes B)$, and $\Sigma^2 := \|(\mathbf{1} \otimes B, \mathbf{1} \otimes B)\|$. Furthermore, define $\Delta := \sup_P\{\|R(P) - P\|\}$, with R the restriction of T to $B(\mathcal{H}) \otimes \mathbf{1}$. Then*

$$\Sigma \geq d(A, \mathcal{Z}) \frac{\frac{1}{2} - \Delta}{\sqrt{\Delta(1 - \Delta)}} \, . \tag{2}$$

This bound is sharp in the sense that for all $\Delta \in [0, \frac{1}{2}]$, there exist T and A for which (2) attains equality.

Proof: For the sharpness, see section 6.5. As for the bound, we may assume $\Delta < \frac{1}{2}$, since inequality (2) is trivially satisfied otherwise. Denote the spectrum of A by $\mathbf{Spec}(A)$. Let $x := \sup(\mathbf{Spec}(A))$ and $y := \inf(\mathbf{Spec}(A))$, so that $d(A, \mathcal{Z}) = \frac{x-y}{2}$. Without loss of generality, assume that there exist normalized eigenvectors ψ_x and ψ_y satisfying $A\psi_x = x\psi_x$ and $A\psi_y = y\psi_y$. (If this

is not the case, choose x' and y' in $\mathbf{Spec}(A)$ arbitrarily close to x and y, and complete the proof using approximate eigenvectors.) Define $\psi_- := \frac{1}{\sqrt{2}}(\psi_x + \psi_y)$, $\tilde{B} := |\psi_-\rangle\langle\psi_-|$ and $\tilde{A} := T(\tilde{B} \otimes \mathbf{1})$.

We thus have $\|[A, \tilde{B}]\| = d(A, \mathcal{Z})$, and furthermore $\|\tilde{A} - \tilde{B}\| \leq \Delta$, so that $\|[A, \tilde{A} - \tilde{B}]\| \leq 2\Delta d(A, \mathcal{Z})$. Then by the triangle inequality $\|[T(\mathbf{1} \otimes B), T(\tilde{B} \otimes \mathbf{1})]\| = \|[A, \tilde{B}] + [A, \tilde{A} - \tilde{B}]\| \geq d(A, \mathcal{Z})(1 - 2\Delta)$, which brings us in a position to apply theorem 3 to the commuting pointers $\tilde{B} \otimes \mathbf{1}$ and $\mathbf{1} \otimes B$. This yields

$$2\Sigma\sqrt{\|(\tilde{B} \otimes \mathbf{1}, \tilde{B} \otimes \mathbf{1})\|} \geq d(A, \mathcal{Z})(1 - 2\Delta). \tag{3}$$

In order to estimate $\|(\tilde{B} \otimes \mathbf{1}, \tilde{B} \otimes \mathbf{1})\|$, we first prove that $\mathbf{Spec}(\tilde{A}) \subseteq [0, \Delta] \cup [(1 - \Delta), 1]$. Let $a \in \mathbf{Spec}(\tilde{A})$. Since T is a contraction and $0 \leq \tilde{B} \leq 1$, we have $0 \leq a \leq 1$. Without loss of generality, assume that there exists a normalized eigenvector ψ_a such that $\tilde{A}\psi_a = a\psi_a$. (Again, if this is not the case, one may use approximate eigenvectors.) Decompose ψ_a over the eigenspaces of $\tilde{B}$, i.e. write $\psi_a = \chi_1 + \chi_0$, with $\chi_1 \perp \chi_0$, $\tilde{B}\chi_1 = \chi_1$ and $\tilde{B}\chi_0 = 0$. Then $(\tilde{A} - \tilde{B})\psi_a = (a - 1)\chi_1 + a\chi_0$. Since $\|\chi_1\|^2 + \|\chi_0\|^2 = 1$, the inequality $\Delta^2 \geq \|(\tilde{A} - \tilde{B})\psi_a\|^2 = (a - 1)^2\|\chi_1\|^2 + a^2\|\chi_0\|^2$ implies that either $|1 - a| \leq \Delta$ or $a \leq \Delta$. We conclude $\mathbf{Spec}(\tilde{A}) \subseteq [0, \Delta] \cup [(1 - \Delta), 1]$, as desired.

This implies $\mathbf{Spec}(\tilde{A} - \tilde{A}^2) \subseteq [0, \Delta(1 - \Delta)]$. Since $\tilde{B}^2 = \tilde{B}$, we may estimate $\|(\tilde{B} \otimes \mathbf{1}, \tilde{B} \otimes \mathbf{1})\| = \|T(\tilde{B} \otimes \mathbf{1}) - T(\tilde{B} \otimes \mathbf{1})^2\| = \|\tilde{A} - \tilde{A}^2\| \leq \Delta(1 - \Delta)$. Combining this with inequality (3) yields $2\Sigma\sqrt{\Delta(1 - \Delta)} \geq d(A, \mathcal{Z})(1 - 2\Delta)$, which was to be demonstrated. $\qquad\square$

In the case of no disturbance, $\Delta = 0$, we see that $\Sigma \to \infty$. No information transfer from $\mathcal{A}$ is allowed if all states on $\mathcal{A}$ are left intact. This is Werner's (see [We2]) formulation of the Heisenberg principle. In the opposite case of perfect information transfer, $\Sigma = 0$, inequality 2 shows that Δ must equal at least one half. We shall see in section 6 that this corresponds with a so-called 'collapse of the wave function'. These two extreme situations are connected by theorem 4 in a continuous fashion, as indicated in the graph below:
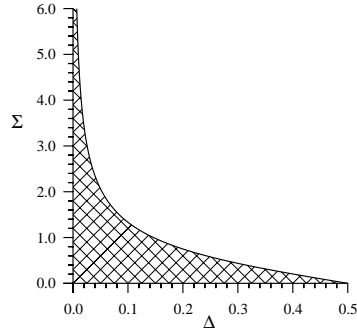

Fig. 2: The combinations (Δ, Σ) below the curve are forbidden, those above are allowed. (With $d(A, \mathcal{Z}) = 1$.)

The upper left corner of the curve illustrates the Heisenberg principle, whereas in the the lower right corner, we can see the collapse of the wave function.

5.2 Heisenberg Principle for General Information Transfer

We now prove a version of the Heisenberg Principle for general information transfer.

Corollary 5 *Let $T : B(\mathcal{H}) \otimes \mathcal{B} \to B(\mathcal{H})$ be a CP-map, let $A \in B(\mathcal{H})$ and $B \in \mathcal{B}$ be Hermitean, $A \notin \mathcal{Z} = \mathbb{C}\mathbf{1}$. Define $\Delta := \sup_P \{\|R(P) - P\|\}$, with R the restriction of T to $B(\mathcal{H}) \otimes \mathbf{1}$. Define $\delta := \sup_S \{\|T(\mathbf{1} \otimes \mathbf{1}_S(B)) - \mathbf{1}_S(A)\|\}$. Then, for δ and Δ in $[0, \frac{1}{2}]$, we have*

$$(\tfrac{1}{2} - \delta)^2 + (\tfrac{1}{2} - \Delta)^2 \leq \tfrac{1}{4} \qquad (4)$$

This bound is sharp in the sense that for all $\Delta \in [0, \frac{1}{2}]$, there exists a T for which (4) attains equality.
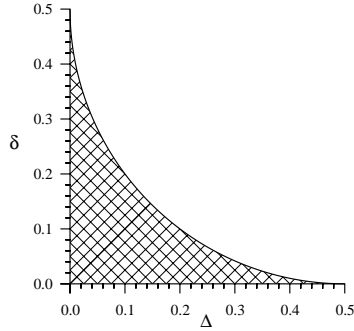

Fig. 3: The combinations (Δ, δ) below the curve are forbidden, those above are allowed.

Proof: Choose a nontrivial subset S of $\mathbf{Spec}(A)$ and put $P := \mathbf{1} \otimes \mathbf{1}_S(B)$. Since $\|T(P) - \mathbf{1}_S(A)\| \leq \delta$ and $\mathbf{Spec}(\mathbf{1}_S(A)) = \{0, 1\}$, we have $\mathbf{Spec}(T(P)) \subseteq [0, \delta] \cup [1 - \delta, 1]$ (cf. the proof of theorem 4). Thus $\Sigma^2 = \|T(P) - T(P)^2\| \leq \delta(1 - \delta)$. Similarly, $d(T(P), \mathcal{Z}) \geq \frac{1}{2} - \delta$ since $\mathbf{Spec}(T(P))$ contains points in both $[0, \delta]$ and $[1 - \delta, 1]$. Apply theorem 4 to the pointer P to obtain $\sqrt{\delta(1 - \delta)} \geq (\frac{1}{2} - \delta)(\frac{1}{2} - \Delta)/\sqrt{\Delta(1 - \Delta)}$, or equivalently $(\frac{1}{2} - \delta)^2 + (\frac{1}{2} - \Delta)^2 \leq \frac{1}{4}$. For sharpness, see section 6.5. $\qquad\qquad\square$

A measurement which does not disturb any state ($\Delta = 0$) cannot yield information ($\delta \geq \frac{1}{2}$). This is the Heisenberg principle. On the other hand, perfect information ($\delta = 0$) implies full disturbance ($\Delta \geq \frac{1}{2}$), corresponding to the collapse of the wave function. Both extremes are connected in a continuous fashion, as depicted above.

5.3 Application: Resonance Fluorescence

Corollary 5 may be used to determine the minimum amount of disturbance if the quality of the measurement is known. Alternatively, if the system is only mildly disturbed, one may find a bound on the attainable measurement quality. Let us concentrate on the latter option.

We investigate the radiation emission of a laser-driven two-level atom. The emitted EM-radiation yields information on the atom. A two-level atom (i.e. a qubit) only has three independent observables: σ_x, σ_y and σ_z. There are various ways to probe the EM field: photon counting, homodyne detection, heterodyne detection, etcetera. For a strong ($\Omega \gg 1$) resonant ($\omega_{\text{laser}} = \omega_{\text{atom}}$) laser, we will use corollary 5 to prove that *any* EM-measurement of σ_x, σ_y or σ_z will have a measurement infidelity of at least

$$\delta \geq \tfrac{1}{2} - \tfrac{1}{2}\sqrt{1 - e^{-\frac{3}{2}\lambda^2 t}},$$

with λ the coupling constant. For a measurement with two outcomes, δ is the maximal probability of getting the wrong outcome.
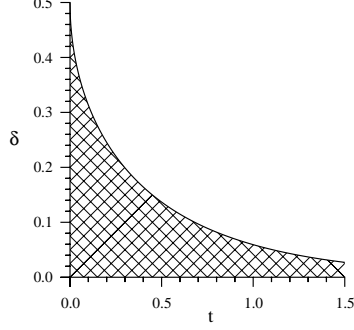
Fig. 4: Lower bound on δ in terms of t (in units of λ^{-2}).

5.3.1 Unitary Evolution on the Closed System

The atom is modelled by the Hilbert-space $\mathbb{C}^2$ (only two energy-levels are deemed relevant). In the field, we discern a *forward* and a *side* channel, each described by a bosonic Fock-space $\mathcal{F}$. The laser is put on the forward channel, which is thus initially in the state ϕ_Ω, the coherent state with frequency ω and strength Ω. (The field strength is parametrized by the frequency of the induced Rabi-oscillations). The side channel starts in the vacuum state ϕ_0. The state at time t is then given by

$$T_t^*(\rho) = U(t)(\rho \otimes \phi_\Omega \otimes \phi_0)U^\dagger(t)\,,$$

with time evolution

$$\frac{d}{dt}U_t = -i(H_S + H_F + \lambda H_I)U_t\,.$$

$H_S \in B(\mathbb{C}^2)$ is the Hamiltonian of the two-level atom, $H_F \in B(\mathcal{F} \otimes \mathcal{F})$ that of the field and $\lambda H_I \in B(\mathbb{C}^2) \otimes B(\mathcal{F} \otimes \mathcal{F})$ is the interaction-Hamiltonian. Define the interaction-picture time evolution by

$$\hat{T}_t^*(\rho) := U_1(t)^\dagger U_2(t)^\dagger T_t^*(\rho) U_2(t) U_1(t)\,,$$

where $U_1(t) := e^{-iH_S t}$ and $U_2(t) := e^{-iH_F t}$ form the 'unperturbed' time evolution.

We now investigate $\hat{T}_t$ instead of T_t. Indeed, we are looking for a bound on the measurement infidelity $\delta = \sup_S\{\|T(\mathbf{1} \otimes \mathbf{1}_S(B)) - \mathbf{1}_S(A)\|\}$ of T. Yet if $\hat{B} := U_2^\dagger B U_2$, then $\hat{T}(\mathbf{1}_S(\hat{B})) = T(\mathbf{1}_S(B))$, so that $\hat{\delta} = \sup_S\{\|\hat{T}(\mathbf{1} \otimes \mathbf{1}_S(\hat{B})) - \mathbf{1}_S(A)\|\} = \delta$. If we find the interaction-picture disturbance $\hat{\Delta}$, corollary 5 will yield a bound on $\hat{\delta}$, and thus on δ.

In the weak coupling limit $\lambda \downarrow 0$, $\hat{T}_t$ is given by $\hat{T}_t^*(\rho) = \hat{U}(t/\lambda^2)(\rho \otimes \phi_\Omega \otimes \phi_0)\hat{U}^\dagger(t/\lambda^2)$, where the evolution of the unitary cocycle $t \mapsto \hat{U}_t$ is described (see [AFL]) by a Quantum Stochastic Differential Equation or QSDE. Explicitly calculating the maximal added variances Σ^2 by solving the QSDE is in general rather nontrivial, if possible at all. (See [JB] for the case of spontaneous decay, i.e. $\Omega = 0$, with the map $\hat{T}_t$ restricted to the commutative algebra of homodyne measurement results.)

5.3.2 Master Equation for the Open System

Fortunately, in contrast to the somewhat complicated time evolution $\hat{T}_t$ of the combined system, the evolution *restricted to the two-level system* is both well-known and uncomplicated. If we use λ^{-2} as a unit of time, then the restricted evolution $\hat{R}_t^*(\rho) := \mathbf{tr}_{\mathcal{F} \otimes \mathcal{F}}\hat{T}_t^*(\rho)$ of the two-level system

is known (see [BGM]) to satisfy the Master equation

$$\frac{d}{dt}\hat{R}_t^*(\rho) = L(\hat{R}_t^*(\rho))\,, \tag{5}$$

with the Liouvillian $L(\rho) := \frac{1}{2}i\Omega[e^{-i(\omega-E)t}V + e^{i(\omega-E)t}V^\dagger, \rho] - \frac{1}{2}\{V^\dagger V, \rho\} + V\rho V^\dagger$. In this expression, E is the energy-spacing of the two-level atom and $V^\dagger = \sigma_+$, $V = \sigma_-$ are its raising and lowering operators. In the case $\omega = E$ of *resonance fluorescence*, we obtain

$$L(\rho) = \tfrac{1}{2}i\Omega[V + V^\dagger, \rho] - \tfrac{1}{2}\{V^\dagger V, \rho\} + V\rho V^\dagger\,.$$

If we parametrize a state by its Bloch-vector $\hat{R}_t^*(\rho) = \frac{1}{2}(\mathbf{1} + x\sigma_x + y\sigma_y + z\sigma_z)$, then equation 5 is simply the following differential equation on $\mathbb{R}^3$:

$$\frac{d}{dt}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{2} & \Omega \\ 0 & -\Omega & -1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

This can be solved explicitly. For $\Omega \gg 1$, the solution approaches

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} e^{-\frac{1}{2}t} & 0 & 0 \\ 0 & e^{-\frac{3}{4}t}\cos(\Omega t) & e^{-\frac{3}{4}t}\sin(\Omega t) \\ 0 & -e^{-\frac{3}{4}t}\sin(\Omega t) & e^{-\frac{3}{4}t}\cos(\Omega t) \end{pmatrix}\begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}\,.$$

If we move to the interaction picture once more to counteract the Rabi-oscillations, i.e. with $U_1(t) = e^{\frac{i}{2}\Omega t\sigma_x}$ and $U_2 = \mathbf{1}$, we see that the time evolution is transformed to

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} e^{-\frac{1}{2}t} & 0 & 0 \\ 0 & e^{-\frac{3}{4}t} & 0 \\ 0 & 0 & e^{-\frac{3}{4}t} \end{pmatrix}\begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}\,.$$

Since the trace distance $D(\rho, \tau)$ is exactly half the Euclidean distance between the Bloch vectors of ρ and τ, (see [NC]), we see that $\Delta = \frac{1}{2}(1 - e^{-\frac{3}{4}t})$. For any measurement of σ_x, σ_y or σ_z, we therefore have $\delta \geq \frac{1}{2} - \frac{1}{2}\sqrt{1 - e^{-\frac{3}{2}t}}$ by corollary 5 (remember that t is in units of λ^{-2}).

6 Collapse of the Wave function

The 'collapse of the wave function' may be seen as the flip side of the Heisenberg principle. It states that if information is extracted from a system, then its states undergo a very specific kind of perturbation, called decoherence.

6.1 Collapse for Unbiased Information Transfer

We start out by investigating unbiased Information Transfer. We prove a sharp upper bound on the amount of remaining coherence in terms of the measurement quality.

Theorem 6 *Let $T : \mathcal{A} \otimes \mathcal{B} \to \mathcal{A}$ be a CP-map. Let $B \in \mathcal{B}$ be Hermitian, and define $A := T(\mathbf{1} \otimes B)$. Suppose that ψ_x and ψ_y are eigenvectors of A with different eigenvalues x and y respectively. Define $R : \mathcal{A} \to \mathcal{A}$ to be the restriction of T to $\mathcal{A} \otimes \mathbf{1}$, and put $\Sigma^2 := \|(B, B)\|$. Then for all $\alpha, \beta \in \mathbb{C}$ with $|\alpha|^2 + |\beta|^2 = 1$, we have*

$$D\left(R^*\left(|\alpha\psi_x + \beta\psi_y\rangle\langle\alpha\psi_x + \beta\psi_y|\right), R^*\left(|\alpha|^2|\psi_x\rangle\langle\psi_x| + |\beta|^2|\psi_y\rangle\langle\psi_y|\right)\right) \leq \frac{\Sigma/|x - y|}{\sqrt{1 + 4\left(\Sigma/|x - y|\right)^2}}\,. \tag{6}$$

11

This bound is sharp in the sense that for all values of $\Sigma/|x-y|$, there exist T, ψ_x, ψ_y, α and β for which (6) attains equality.

Proof: The l.h.s. of (6) equals $\sup\{\bar{\alpha}\beta\langle\psi_x, R(P)\psi_y\rangle + \mathbf{c.c.} \mid P \in \mathcal{A}, P^2 = P^\dagger = P\}$. Furthermore, $2|\alpha||\beta| \leq 1$, so that it suffices to bound the 'coherence' $\langle\psi_x, R(P)\psi_y\rangle$ on all projections P. Now $(x-y)\langle\psi_x, R(P)\psi_y\rangle = \langle\psi_x, [A, R(P)]\psi_y\rangle$, and $[A, R(P)] = (P \otimes \mathbf{1}, \mathbf{1} \otimes B) - (\mathbf{1} \otimes B, P \otimes \mathbf{1})$. Thus

$$(x-y)\langle\psi_x, R(P)\psi_y\rangle = \langle\psi_x, (P \otimes \mathbf{1}, \mathbf{1} \otimes B)\psi_y\rangle - \langle\psi_x, (\mathbf{1} \otimes B, P \otimes \mathbf{1})\psi_y\rangle, \qquad (7)$$

and we will bound these last two terms. In the notation of lemma 1, we have $\|g(\mathbf{1} \otimes B)\| = \Sigma$. Therefore $\langle\psi_x, (P \otimes \mathbf{1}, \mathbf{1} \otimes B)\psi_y\rangle = \langle g(P \otimes \mathbf{1})\psi_x, g(\mathbf{1} \otimes B)\psi_y\rangle \leq \|g(P \otimes \mathbf{1})\psi_x\|\|g(\mathbf{1} \otimes B)\psi_y\| \leq \Sigma\sqrt{\langle\psi_x, (P \otimes \mathbf{1}, P \otimes \mathbf{1})\psi_x\rangle}$. We will bound $\langle\psi_x, (T(P^2 \otimes \mathbf{1}) - T(P \otimes \mathbf{1})^2)\psi_x\rangle = \langle\psi_x, (R(P) - R(P)^2)\psi_x\rangle$ in terms of the coherence. For brevity, denote $X_{xx'} := \langle\psi_x, X\psi_{x'}\rangle$. Since $\psi_x \perp \psi_y$, we have $(R(P)^2)_{xx} \geq |R(P)_{xx}|^2 + |R(P)_{xy}|^2$, so that $(R(P) - R(P)^2)_{xx} \leq R(P)_{xx}(1 - R(P)_{xx}) - |R(P)_{xy}|^2$. Since $x(1-x) \leq \frac{1}{4}$ for all $x \in \mathbb{R}$, this is at most $\frac{1}{4} - |R(P)_{xy}|^2$. All in all, we have obtained $(P \otimes \mathbf{1}, \mathbf{1} \otimes B)_{xy} \leq \Sigma\sqrt{\frac{1}{4} - |R(P)_{xy}|^2}$, and of course the same for $x \leftrightarrow y$. Plugging these into equation 7 yields $|x - y||R(P)_{xy}| \leq 2\Sigma\sqrt{\frac{1}{4} - |R(P)_{xy}|^2}$, or equivalently $|R(P)_{xy}| \leq \frac{\Sigma/|x-y|}{\sqrt{1+4(\Sigma/|x-y|)^2}}$, which was to be proven. For sharpness, see section 6.5. $\qquad\square$

Consider the ideal case of perfect ($\Sigma = 0$) information transfer. Suppose that the system $\mathcal{A}$ is initially in the coherent state $|\alpha\psi_x + \beta\psi_y\rangle\langle\alpha\psi_x + \beta\psi_y|$. Then theorem 6 says that, after the information transfer to the ancilla $\mathcal{B}$, the system $\mathcal{A}$ cannot be distinguished from one that started out in the 'incoherent' state $|\alpha|^2|\psi_x\rangle\langle\psi_x| + |\beta|^2|\psi_y\rangle\langle\psi_y|$. As far as the behaviour of $\mathcal{A}$ is concerned, it is therefore completely harmless to assume that a collapse $|\alpha\psi_x + \beta\psi_y\rangle\langle\alpha\psi_x + \beta\psi_y| \mapsto |\alpha|^2|\psi_x\rangle\langle\psi_x| + |\beta|^2|\psi_y\rangle\langle\psi_y|$ has occurred at the start of the procedure.

Now consider the other extreme of a measurement which leaves all states intact, i.e. $R^*(\rho) = \rho$ for all ρ. Then there exist states for which the l.h.s. of equation (6) equals $\frac{1}{2}$, forcing $\Sigma \to \infty$; no information can be obtained. This is Werner's formulation of the Heisenberg principle.

Theorem 6 thus unifies the Heisenberg principle and the collapse of the wave function. For $\Sigma = 0$ we have a full decoherence, whereas if all states are left intact, we have $\Sigma \to \infty$. For all intermediate cases, the bound 6 on the remaining coherence is an increasing function of $\Sigma/|x-y|$.
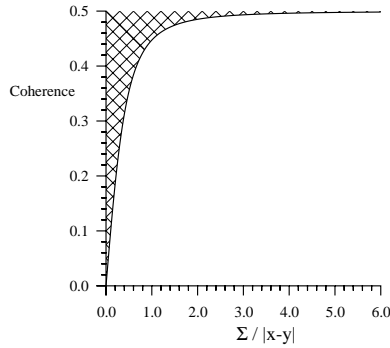

Fig. 5: Bound on the coherence as a function of $\Sigma/|x-y|$. All points above this curve are forbidden, all points below are allowed.

This agrees with physical intuition: decoherence between ψ_x and ψ_y is expected to occur in case the information transfer is able to distinguish between the two. This is the case if the variance is small w.r.t. the differences in mean.

6.2 Application: Perfect Qubit Measurement

In section 2, we have encountered the von Neumann Qubit measurement. Now consider any perfect measurement T of σ_z with pointer $\mathbf{1} \otimes (\delta_+ - \delta_-)$ which leaves $|\uparrow\rangle\langle\uparrow|$ and $|\downarrow\rangle\langle\downarrow|$ in place, i.e. $R^*(|\uparrow\rangle\langle\uparrow|) = |\uparrow\rangle\langle\uparrow|$ and $R^*(|\downarrow\rangle\langle\downarrow|) = |\downarrow\rangle\langle\downarrow|$. (Such a measurement is called nondestructive.) Theorem 6 then reads $R^*(|\alpha \uparrow + \beta \downarrow\rangle\langle\alpha \uparrow + \beta \downarrow|) = |\alpha|^2|\uparrow\rangle\langle\uparrow| + |\beta|^2|\downarrow\rangle\langle\downarrow|$, illustrated below.
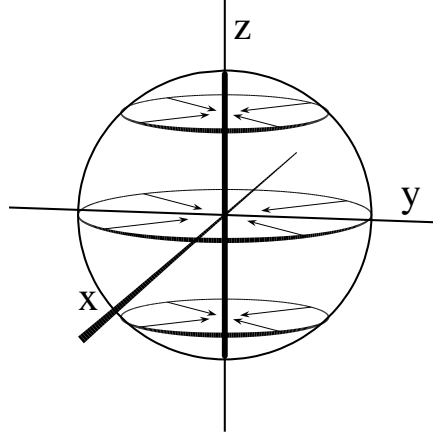

Fig. 6: Collapse on the Bloch-sphere for perfect measurement.

Incidentally, the trace distance between the centre of the Bloch sphere and its surface is $\frac{1}{2}$, so that we read off $\Delta = \sup\{D(R^*(\rho), \rho); \rho \in \mathcal{S}(M_2)\} = \frac{1}{2}$. This was predicted by theorem 4.

6.3 Collapse of the Wave function for General Measurement

We will prove a sharp bound on the remaining coherence in general information transfer. For technical convenience, we will focus attention on nondestructive measurements. A measurement of A is called 'nondestructive' (or 'conserving' or 'quantum nondemolition') if it leaves the eigenstates of A intact, so that repetition of the measurement will yield the same result. For example, the measurement in section 2.2 is nondestructive, the one in section 5.3 is destructive. Restriction to nondestructive measurements is quite common in quantum measurement theory (see [Per]).

Corollary 7 *Let $T : B(\mathcal{H}) \otimes \mathcal{B} \to B(\mathcal{H})$ be a CP-map, let $A \in B(\mathcal{H})$ and $B \in \mathcal{B}$ be Hermitean and let $\{\psi_i\}$ be an orthogonal basis of eigenvectors of A, with eigenvalues a_i. Define the measurement infidelity $\delta := \sup_S\{\|T(\mathbf{1} \otimes \mathbf{1}_S(B)) - \mathbf{1}_S(A)\|\}$. Suppose that T is nondestructive, i.e. $R^*(|\psi_i\rangle\langle\psi_i|) = |\psi_i\rangle\langle\psi_i|$ for all ψ_i, with R the restriction of T to $B(\mathcal{H}) \otimes \mathbf{1}$. Then if $\delta \in [0, \frac{1}{2}]$, and $a_i \neq a_j$,*

$$D\left(R^*\left(|\alpha\psi_i + \beta\psi_j\rangle\langle\alpha\psi_i + \beta\psi_j|\right), \left(|\alpha|^2|\psi_i\rangle\langle\psi_i| + |\beta|^2|\psi_j\rangle\langle\psi_j|\right)\right) \leq \sqrt{\delta(1-\delta)}. \quad (8)$$

This bound is sharp in the sense that for all $\delta \in [0, \frac{1}{2}]$, there exist T, ψ_i, ψ_j, α and β for which (8) attains equality.
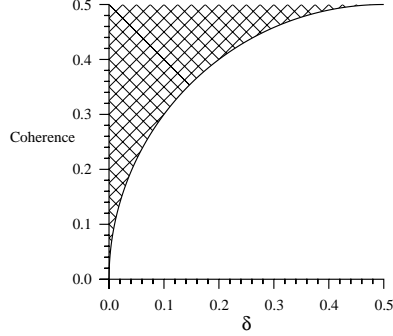
13

Fig. 7: Bound on the coherence in terms of δ. All points above the curve are forbidden, all points below are allowed.

Proof: For sharpness, see section 6.5. Choose a set S such that $a_i \in S$ and $a_j \notin S$. T is an unbiased measurement of $T(\mathbf{1} \otimes \mathbf{1}_S(B))$ with pointer $\mathbf{1}_S(B)$ and maximal added variance $\Sigma^2 \leq \delta(1-\delta)$ (cf. the proof of corollary 5). We will prove that ψ_i and ψ_j are eigenvectors of $T(\mathbf{1} \otimes \mathbf{1}_S(B))$ with eigenvalues x and y which differ at least $1-2\delta$.

Define $P_i := |\psi_i\rangle\langle\psi_i|$. Since T is nondestructive, we have $\langle\psi_j, R(P_i)\psi_j\rangle = \langle\psi_j, P_i\,\psi_j\rangle$ for all j. Apparently, $R(P_i)$ has only one nonzero diagonal element, a 1 at position (i,i). Since $R(P_i) \geq 0$, this implies $R(P_i) = P_i$. Then $(P_i \otimes \mathbf{1}, P_i \otimes \mathbf{1}) = 0$, so that by Cauchy-Schwarz $(P_i \otimes \mathbf{1}, \mathbf{1} \otimes \mathbf{1}_S(B)) = 0$. Since $P_i = T(P_i \otimes \mathbf{1})$, we have $[T(\mathbf{1} \otimes \mathbf{1}_S(B)), P_i] = (P_i \otimes \mathbf{1}, \mathbf{1} \otimes \mathbf{1}_S(B)) - (\mathbf{1} \otimes \mathbf{1}_S(B), P_i \otimes \mathbf{1}) = 0$. Therefore ψ_i is an eigenvector of $T(\mathbf{1} \otimes \mathbf{1}_S(B))$, with eigenvalue x, say. By a similar reasoning, ψ_j is also an eigenvector, denote its eigenvalue by y.

Since $\|T(\mathbf{1} \otimes \mathbf{1}_S(B)) - \mathbf{1}_S(A)\| \leq \delta$, we have in particular $\|(T(\mathbf{1} \otimes \mathbf{1}_S(B)) - \mathbf{1}_S(A))\psi_i\| = |x - 1| \leq \delta$ and $\|(T(\mathbf{1} \otimes \mathbf{1}_S(B)) - \mathbf{1}_S(A))\psi_j\| = |y| \leq \delta$, so that $|x - y| \geq 1 - 2\delta$. We can now apply Theorem 6. On the l.h.s. of the bound (6), we may substitute $R^*(|\alpha|^2|\psi_i\rangle\langle\psi_i| + |\beta|^2|\psi_j\rangle\langle\psi_j|) = |\alpha|^2|\psi_i\rangle\langle\psi_i| + |\beta|^2|\psi_j\rangle\langle\psi_j|$ on account of T being nondestructive. On the r.h.s., we substitute $\Sigma = \sqrt{\delta(1-\delta)}$ and $|x - y| = (1 - 2\delta)$. Strikingly enough, this yields the bound

$$\left(\sqrt{\delta(1-\delta)}/(1-2\delta)\right)/\sqrt{1 + 4\left(\delta(1-\delta)/(1-2\delta)^2\right)} = \sqrt{\delta(1-\delta)}. \qquad \square$$

For perfect measurement ($\delta = 0$), this yields $R^*(|\alpha\psi_i + \beta\psi_j\rangle\langle\alpha\psi_i + \beta\psi_j|) = |\alpha|^2|\psi_i\rangle\langle\psi_i| + |\beta|^2|\psi_j\rangle\langle\psi_j|$; all coherence between ψ_i and ψ_j must vanish. This collapse of the wave function is illustrated in the lower left corner of fig. 7. On the other hand, if all states are left intact so that $R^* = Id$, then we must have $\delta = \frac{1}{2}$; no information can be gained. This is illustrated in the upper right corner of fig.7. Corollary 7 is a unified description of the Heisenberg principle and the collapse of the wave function.

6.4   Application: Nondestructive Qubit-Measurement

In quantum information theory, a σ_z-measurement is often taken to yield output $+1$ or -1, according to whether the input was $|\uparrow\rangle$ or $|\downarrow\rangle$. It is nondestructive if it leaves the states $|\uparrow\rangle$ and $|\downarrow\rangle$ intact, yet it is only unbiased if it is perfect. Corollary 7 shows that in the nondestructive case, the Bloch-sphere collapses to the cigar-shaped region depicted below:
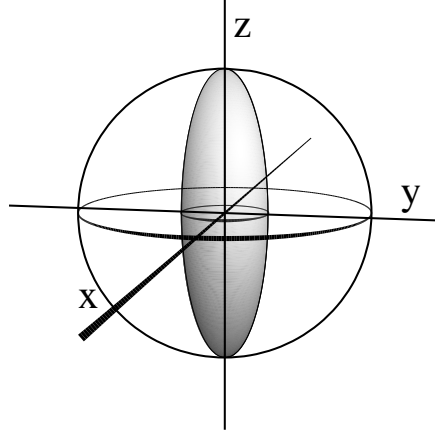
14

Fig. 8: Collapse on the Bloch-sphere with $\delta = 0.01$.

Current single-qubit readout technology is just now moving into the regime where the bound (8) becomes significant. in [Lu], a nondestructive measurement of a SQUID-qubit was described, with experimentally determined measurement infidelity $\delta = 0.13$. The bound then equals 0.336.

6.5 Sharpness of the Bounds

We have yet to prove sharpness of all bounds. Let $V_+ := \begin{pmatrix} \sqrt{1-p} & 0 \\ 0 & \sqrt{p} \end{pmatrix}$, $V_- := \begin{pmatrix} \sqrt{p} & 0 \\ 0 & \sqrt{1-p} \end{pmatrix}$, and define $T : M_2 \otimes \mathcal{C}(\Omega) \to M_2$ by $T(X \otimes f) := f(+1)V_+ X V_+ + f(-1)V_- X V_-$. For $p = 0$, this is the von Neumann-measurement. As a measurement of σ_z with pointer $B := (\delta_+ - \delta_-)/(1-2p)$, we have $\delta = p$. This yields bounds on the disturbance and on the coherence. Corollary 5 and theorem 4 yield $\Delta \geq \frac{1}{2} - \sqrt{p(1-p)}$, corollary 7 and theorem 6 yield $D\Big(R^*\big(|\alpha \uparrow + \beta \downarrow\rangle\langle\alpha \uparrow + \beta \downarrow|\big),$
$\big(|\alpha|^2 |\uparrow\rangle\langle\uparrow| + |\beta|^2 |\downarrow\rangle\langle\downarrow|\big)\Big) \leq \sqrt{p(1-p)}$. We now explicitly calculate the restriction of T to M_2, and find

$$R^*(\rho) = \begin{pmatrix} \rho_{11} & 2\sqrt{p(1-p)}\rho_{12} \\ 2\sqrt{p(1-p)}\rho_{21} & \rho_{22} \end{pmatrix} .$$

The maximal remaining coherence occurs for $\alpha = \beta = 1/\sqrt{2}$, for which it equals $\sqrt{p(1-p)}$. The maximal disturbance equals $\Delta = \frac{1}{2} - \sqrt{p(1-p)}$. This shows all bounds to be sharp.

7 Conclusion

Our investigation of joint measurement, the Heisenberg principle and decoherence has yielded the following results.

I Theorem 3 provides a sharp, state independent bound on the performance of unbiased joint measurement of noncommuting observables. In the case of perfect ($\Sigma = 0$) measurement of one observable, it implies that no information whatsoever ($\Sigma' = \infty$) can be gained on the other.

II Theorem 4 (for unbiased information transfer) and corollary 5 (for general information transfer) provide a sharp, state independent bound on the performance of a measurement in terms of the maximal disturbance that it causes. In the case of zero disturbance, when all states are left intact, it follows that no information can be obtained. This is the Heisenberg principle.

III Theorem 6 (for unbiased information transfer) and corollary 7 (for general information transfer) provide a sharp upper bound on the amount of coherence which can survive information transfer. For perfect information transfer, all coherence vanishes. This clearly proves that decoherence on a system is a mathematical consequence of information transfer out of this system. If, on the other hand, all states are left intact, then it follows that no information can be obtained. This is the Heisenberg principle. Theorem 6 and corollary 7 connect these two extremes in a continuous fashion; they form a unified description of the Heisenberg principle and the collapse of the wave function.

Acknowledgements

References

[AFL] L. Accardi, A. Frigerio, Y.G. Lu, *'The Weak Coupling Limit as a Quantum Functional Central Limit'*, Commun. Math. Phys.**131**, 537–570, (1990).

[AK] E. Arthurs, J. Kelly, *'On Simultaneous Measurement on a Pair of Conjugate Observables'*, Bell. Syst. Tech. J. **44**, 725, (1965).

[BGM] L. Bouten, M. Guţă, H. Maassen, *'Stochastic Schrödinger Equations'*, J. Phys. A. **37**, 3189–3209, (2004).

[Ha] M. Hall, *'Prior Information: How to circumvent the Standard Joint-Measurement Uncertainty Relation'*, Phys. Rev. A **69**, 052113, (2004).

[He] W. Heisenberg, *'Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik'*, Z. Phys. **43**, 172–198, (1927).

[Hp] K. Hepp, *'Quantum Theory of Measurement and Macroscopic Observables'*, Helv. Phys. Acta **45**, 237–248, (1972).

[Ho] A. S. Holevo, *'Probabilistic and Statistical Aspects of Quantum Theory'*, North Holland Publishing Company, (1982).

[Is] S. Ishikawa, *'Uncertainty Relations in Simultaneous Measurements for Arbitrary Observables'*, Rep. Math. Phys. **29**, 257–273, (1991).

[JB] B. Janssens, L. Bouten, *'Optimal Pointers for Joint Measurement of σ_x and σ_z via Homodyne Detection'*, J. Phys. A. **39**, 2773–2790, (2006).

[JM] B. Janssens, H. Maassen, *'Information Transfer Implies State Collapse'*, `www.arxiv.org/abs/quant-ph/0602140`, (2006).

[JZ] E. Joos, H. Zeh, *'The Emergence of Classical Properties Through Interaction with the Environment'*, Z. Phys. B **59**, 223–243, (1985).

[Ke] E. Kennard, *'Zur Quantenmechanik einfacher Bewegungstypen'*, Z. Phys. **44**, 326–325, (1927).

[Lu] A. Lupaşcu e.a. , *'High-Contast Dispersive Readout of a Superconducting Flux Qubit using a Nonlinear Resonator'*, Phys. Rev. Lett. **96**, 127003, (2006).

[Ne]   J. von Neumann, *'Mathematische Grundlagen der Quantenmechanik'*, Springer-Verlag, (1932).

[NC]   M. Nielsen, I. Chuang, *'Quantum Computation and Quantum Information'*, Cambridge University Press, (2000).

[Oz]   M. Ozawa, *'Universally valid Reformulation of the Heisenberg Uncertainty Principle on Noise and Disturbance in Measurement'*, Phys. Rev. A **67**, 042105 (2003).

[Per]   A. Peres, *'Quantum Theory: Concepts and Methods'*, Kluwer Academic Publishers (1993).

[Ro]   H. Robertson, *'The Uncertainty Principle'*, Phys. Rev. **34**, 163–164, (1929).

[Se]   G. Sewell, *'On the Mathematical Structure of Quantum Measurement Theory'*, Rep. Math. Phys. **56**, 271–290, (2005).

[Ta]   M. Takesaki, *'Theory of Operator Algebras I'*, Springer-Verlag, New York, (1979).

[We1]   R. Werner, *'Optimal Cloning of Pure States'*, Phys. Rev. A, **58**, 1827–1832, (1998).

[We2]   R. Werner, *'Quantum Information Theory – an Invitation'*, Springer Tracts in Modern Physics **173**, 14–57, (2001).

[Zu]   W. Zurek, *'Environment-Induced Superselection Rules'*, Phys. Rev. D **26**, 1862–1880, (1982).